

A Detection of the Baryon Acoustic Oscillation features in the SDSS BOSS DR12 Galaxy Bispectrum

David W. Pearson^{1★} and Lado Samushia^{1,2,3}

¹*Department of Physics, Kansas State University, 116 Cardwell Hall, Manhattan, KS, 66506, USA*

²*National Abastumani Astrophysical Observatory, Ilia State University, 2A Kazbegi Ave., GE-1060 Tbilisi, Georgia*

³*Institute of Cosmology & Gravitation, University of Portsmouth, PO1 3FX, UK*

Accepted 2018 May 8. Received 2018 May 3; in original form 2017 December 13

ABSTRACT

We present the first high-significance detection (4.1σ) of the Baryon Acoustic Oscillations (BAO) feature in the galaxy bispectrum of the twelfth data release (DR12) of the Baryon Oscillation Spectroscopic Survey (BOSS) constant mass sample ($0.43 \leq z \leq 0.7$). We measured the scale dilation parameter, α , using the power spectrum, bispectrum, and both simultaneously for DR12, plus 2048 MultiDark-PATCHY mocks in the North and South Galactic Caps (NGC and SGC, respectively), and the volume-weighted averages of those two samples (N+SGC). The fitting to the mocks validated our analysis pipeline, yielding values consistent with the mock cosmology. By fitting to the power spectrum and bispectrum separately, we tested the robustness of our results, finding consistent values from the NGC, SGC, and N+SGC in all cases. We found $D_V = 2032 \pm 24(\text{stat.}) \pm 15(\text{sys.})$ Mpc, $D_V = 2038 \pm 55(\text{stat.}) \pm 15(\text{sys.})$ Mpc, and $D_V = 2031 \pm 22(\text{stat.}) \pm 10(\text{sys.})$ Mpc from the N+SGC power spectrum, bispectrum, and simultaneous fitting, respectively. Our bispectrum measurement precision was mainly limited by the size of the covariance matrix. Based on the fits to the mocks, we showed that if a less noisy estimator of the covariance were available, from either a theoretical computation or a larger suite of mocks, the constraints from the bispectrum and simultaneous fits would improve to 1.1 per cent (1.3 per cent with systematics) and 0.7 per cent (0.9 per cent with systematics), respectively, with the latter being slightly more precise than the power spectrum only constraints from the reconstructed field.

Key words: cosmology: observations – distance scale – large-scale structure of Universe.

1 INTRODUCTION

The bispectrum is sensitive to non-Gaussianities in the galaxy density field from primordial physics, gravitational dynamics, velocity distortions, and biasing. However, bispectrum estimates are quite noisy since one can only average over triangles of the same shape, but with different orientations. This means that the number of coherence cells that contribute to a single bispectrum measurement is relatively small (Martínez & Saar 2002). As the volume of our surveys increases, the noise in the bispectrum should decrease, making it a potentially powerful tool in improving constraints on cosmological parameters [see Sefusatti & Scoccimarro 2005, equations (18) & (19)]. In particular, future surveys such as the Wide Field InfraRed Survey Telescope (WFIRST; Spergel et al. 2015) surveys and the Dark Energy Spectroscopic Instrument (DESI) Bright Galaxy Survey (DESI Collaboration et al. 2016), will simultaneously cover a large volume and have a high number density, making the constraining power of the bispectrum comparable to that of the power spectrum (Gagrani & Samushia 2017).

Recent studies, using the Sloan Digital Sky Survey’s (SDSS) Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al. 2013) data, have used the galaxy bispectrum to help bolster constraints of various cosmological parameters. Gil-Marín et al. (2015) used the galaxy power spectrum and bispectrum monopoles to constrain a model with two bias parameters, b_1 and b_2 , for a non-linear, non-local bias model, along with the linear growth parameter, f . It was the additional constraining power of the bispectrum that allowed them to break the degeneracy between the bias and growth. More recently, Gil-Marín et al. (2017) used the galaxy power spectrum monopole and quadrupole along with the bispectrum monopole to perform a measurement of redshift-space distortions.

The Baryon Acoustic Oscillation (BAO) peak position allows us to constrain the expansion rate of our Universe by measuring its size at a number of different redshifts. The measurement, in its simplest form, involves constraining a single scale dilation parameter, α , which can then be related to the distance to a particular redshift

$$\alpha \equiv \frac{D_V(z)r_d^{\text{fid}}}{D_V^{\text{fid}}(z)r_d}, \quad (1)$$

★ E-mail: dpearson@phys.ksu.edu

where r_d is the sound horizon at the drag epoch, and the superscript ‘fid’ refers to the fiducial cosmology. That is, the distance to redshift z , is

$$D_V(z) = \alpha D_V^{\text{fid}}(z) \left(\frac{r_d}{r_d^{\text{fid}}} \right). \quad (2)$$

This distance is also related to the angular diameter distance, and the Hubble parameter at the redshift of interest via

$$D_V(z) = [cz(1+z)^2 D_A^2(z) H^{-1}(z)]^{1/3}. \quad (3)$$

We can compute $H(z)$ simply as $H(z) = H_0 E(z)$, with (Peebles 1980; Martínez & Saar 2002)

$$E^2(z) = \Omega_M(1+z)^3 + \Omega_\Lambda - \frac{kc^2}{H_0^2 R_0^2} (1+z)^2, \quad (4)$$

and $k = 0$ for a spatially flat cosmology. The angular diameter distance is given, in general, by

$$D_A(z) = \frac{R_0}{1+z} S_k[\omega(z)], \quad (5)$$

where R_0 is the radius of curvature of the Universe, and

$$S_k[\omega(z)] = \begin{cases} \sin \omega(z), & k = 1, \\ \omega(z), & k = 0, \\ \sinh \omega(z), & k = -1. \end{cases} \quad (6)$$

So, in a flat universe, $S_k[\omega(z)]$ is simply $\omega(z)$, which is given by (see Martínez & Saar 2002, chapter 2)

$$\omega(z) = \frac{c}{R_0 H_0} \int_0^z \frac{dz'}{E(z')}, \quad (7)$$

giving the angular diameter distance as

$$D_A(z) = \frac{c}{(1+z)H_0} \int_0^z \frac{dz'}{E(z')}. \quad (8)$$

From this, given a fiducial cosmology and a constraint on α , we can easily calculate the distance, $D_V(z)$.

When it comes to the BAO peak position, studies tend to focus on two-point statistics – the correlation function and its Fourier transform, the power spectrum (see e.g. Anderson et al. 2012, 2014; Cuesta et al. 2016; Gil-Marín et al. 2016; Beutler et al. 2017; Ross et al. 2017). This is justified by the fact that the reconstruction of the galaxy field (Eisenstein et al. 2007) is believed to at least partially ‘move’ the information from the three-point statistics – e.g. the three-point correlation function or the bispectrum – into the two-point statistics. However, measuring the BAO peak position from the bispectrum directly – or the combination of the non-reconstructed power spectrum and bispectrum – is still useful as the effect of the reconstruction on the information content of the higher order statistics is, at the moment, not completely clear.

In a series of recent papers (Slepian & Eisenstein 2015; Slepian et al. 2017a,b) it was shown that the BAO peak could be measured in the galaxy three-point correlation function. Specifically, Slepian et al. (2017b), using the SDSS BOSS Data Release 12 (DR12) constant mass (CMASS) sample, were able to obtain a distance to redshift $z=0.57$ with an accuracy of 1.7 per cent. Additionally, they conclude that the three-point correlation function contains significant additional information on the distance scale.

Inspired by the results of Slepian et al. (2017b), we set out to investigate whether the same signal could be detected in the galaxy bispectrum. In this paper, we present constraints on the scale dilation parameter, α , from the SDSS BOSS DR12 CMASS power spectrum, bispectrum, and a simultaneous constraint from their combination.

We did this separately for the north galactic cap (NGC), south galactic cap (SGC), and volume-weighted north plus south galactic cap (N+SGC) samples for both the data and the mock galaxy catalogues used. The mock results were used as validation for our analysis pipeline and assessing systematic errors. We compared our results with the analyses of Anderson et al. (2014), Cuesta et al. (2016), Gil-Marín et al. (2016), Ross et al. (2017), and Slepian et al. (2017b), finding them to be consistent.

Our constraint on α from the bispectrum only fitting is weaker compared to the results presented in Slepian et al. (2017b). This is chiefly due to the number of mocks used for the covariance matrix estimation, being only about three times the number of data points in the bispectrum measurement, which inflated our uncertainty in the parameter. We show that if this additional uncertainty due to the noisy covariance was negligible, our constraints on the distance scale would be comparable to the ones derived in Slepian et al. (2017b). When combined with the power spectrum measurements the resulting constraint on α reached one per cent and is therefore comparable with the constraints from the power spectrum only of the reconstructed galaxy field presented in Cuesta et al. (2016). If not for the covariance matrix noise, the joint power spectrum and bispectrum constraints would be slightly superior.

The layout of the paper is as follows: In Section 2 we briefly discuss the data and mocks that were used in our analysis, and the procedures used to measure the power spectrum, bispectrum, and covariance. In Section 3 we present the models that were fitted to the data and outline our fitting procedure. In Section 4 we present the main results of this paper, and then provide some discussion in Section 5.

Our fiducial cosmology was chosen to match the *Planck* cosmic microwave background results (Planck Collaboration et al. 2016), with $\Omega_{M,0} = 0.3089$, $\Omega_{\Lambda,0} = 0.6911$, $\Omega_b h^2 = 0.02230$, $h = 0.6774$, $n_s = 0.9667$, and $\sigma_8 = 0.8159$ (i.e. all values were taken from the last column of table 4 in Planck Collaboration et al. 2016).

2 DATA, MOCKS, AND MEASUREMENTS

We used the BOSS DR12 CMASS sample containing 777 202 galaxies in the redshift range $0.43 \leq z \leq 0.7$ (see Alam et al. 2017, for more details about this sample). This sample contains luminous red galaxies selected so as to ensure a roughly constant mass of all tracers, split in two volumes, the NGC and the SGC. We calculated the effective volumes as $1.64h^{-3}\text{Gpc}^3$ and $0.606h^{-3}\text{Gpc}^3$ for the NGC and SGC, respectively, corresponding to effective fundamental frequencies of $0.0053h\text{Mpc}^{-1}$ and $0.0074h\text{Mpc}^{-1}$. For the purposes of covariance estimation and analysis pipeline validation, we used the 2048 MultiDark-PATCHY mock catalogues (Kitaura et al. 2016; Rodríguez-Torres et al. 2016) released with DR12.

2.1 Measuring the power spectrum

To measure the galaxy power spectrum monopole, we used the standard Feldman, Kaiser & Peacock (1994, hereafter FKP) method. To start with, we calculated the weight for each galaxy using the provided FKP and systematic weights via the scheme suggested by Anderson et al. (2012),

$$w(\mathbf{r}) = w_{\text{FKP}}(\mathbf{r}) w_{\text{sys}}(\mathbf{r}) (w_{\text{rf}} + w_{\text{cp}} - 1), \quad (9)$$

where w_{FKP} are the usual FKP weights, w_{sys} are the combined systematics, w_{rf} and w_{cp} are the redshift failure and close pair weights, respectively, which are unity by default. Next we binned the galaxies and randoms onto a $512 \times 1024 \times 512$ grid, in a box of size

$(1792 \times 3584 \times 1792)h^{-3}\text{Mpc}^3$ using a cloud-in-cell (CIC) interpolation scheme and the weights to give $n_{\text{gal}}(\mathbf{r})$ and $n_{\text{ran}}(\mathbf{r})$, respectively. Then we calculated the overdensity field

$$\delta(\mathbf{r}) = n_{\text{gal}}(\mathbf{r}) - \alpha_{\text{den}} n_{\text{ran}}(\mathbf{r}), \quad (10)$$

where α_{den} is the ratio of the sum of the galaxy weights to the sum of the random weights. This was then Fourier transformed using the Fastest Fourier Transform in the West (FFTW) library,¹ and then averaged by frequency in bins of width $\Delta k = 0.008h\text{Mpc}^{-1}$,

$$\hat{P}(k) = \frac{1}{\int \bar{n}^2(\mathbf{r})w^2(\mathbf{r})d\mathbf{r}} [\langle |\delta(\mathbf{k})|^2 \rangle - S_P] G^2(k). \quad (11)$$

Subtracting the shot noise before applying the CIC grid correction is not technically correct, however this is what was done in this work. We note that this has no effect on our results as including the term proportional to k^2 in our power spectrum smoothing polynomial, see equation (27), accounts for the distortion in the broad-band shape caused by this mistake. In equation (11), $\bar{n}(\mathbf{r})$ is the average number density at \mathbf{r} , S_P is the shot noise defined by

$$S_P = \int_{\text{gal}} \bar{n}(\mathbf{r})w^2(\mathbf{r})d\mathbf{r} + \alpha_{\text{den}} \int_{\text{ran}} \bar{n}(\mathbf{r})w^2(\mathbf{r})d\mathbf{r}, \quad (12)$$

and $G(k)$ is the CIC binning correction (Jeong 2010),

$$G(\mathbf{k}) = \left[\text{sinc}\left(\frac{k_x L_x}{N_x}\right) \text{sinc}\left(\frac{k_y L_y}{N_y}\right) \text{sinc}\left(\frac{k_z L_z}{N_z}\right) \right]^{-2}, \quad (13)$$

where $\text{sinc}(x) = \sin(x)/x$, L_i is the length of the box you placed your sample into in the i direction, and N_i is the number of grid cells in that dimension. We did not use a reconstruction procedure (Eisenstein et al. 2007) on our overdensity field, as this further correlates the power spectrum and bispectrum measurements (see Slepian et al. 2017b, section 8.2) and removes most of the bispectrum signal. When combining the NGC and SGC, because of the large size of the combined power spectrum and bispectrum covariance matrix, we used a simple volume-weighted averaging instead of a potentially more efficient inverse covariance matrix-based weighting.

First, we calculated the effective survey volumes of the NGC and SGC via (Anderson et al. 2012),

$$V_{\text{eff}} = \sum_i \left(\frac{\bar{n}(z_i)P_0}{1 + \bar{n}(z_i)P_0} \right)^2 \Delta V(z_i), \quad (14)$$

where P_0 is the approximate power spectrum amplitude that the FKP weights were optimizing for, $\bar{n}(z_i)$ is the average number density of galaxies in the redshift bin z_i , and $\Delta V(z_i)$ is the volume of the spherical shell at that redshift scaled by the fraction of the sky covered by the survey. The weights were then

$$w_{V,i} = \frac{V_{\text{eff},i}}{V_{\text{eff,NGC}} + V_{\text{eff,SGC}}}, \quad (15)$$

where i was either NGC or SGC. The combined power spectrum measurement was then

$$\hat{P}_{\text{N+SGC}}(k) = w_{V,\text{NGC}} \hat{P}_{\text{NGC}}(k) + w_{V,\text{SGC}} \hat{P}_{\text{SGC}}(k). \quad (16)$$

The measured power spectra are shown in Fig. 1, along with the measurements normalized by equation (27) calculated with the best-fitting parameter values. All the measurements agree with each other quite well.

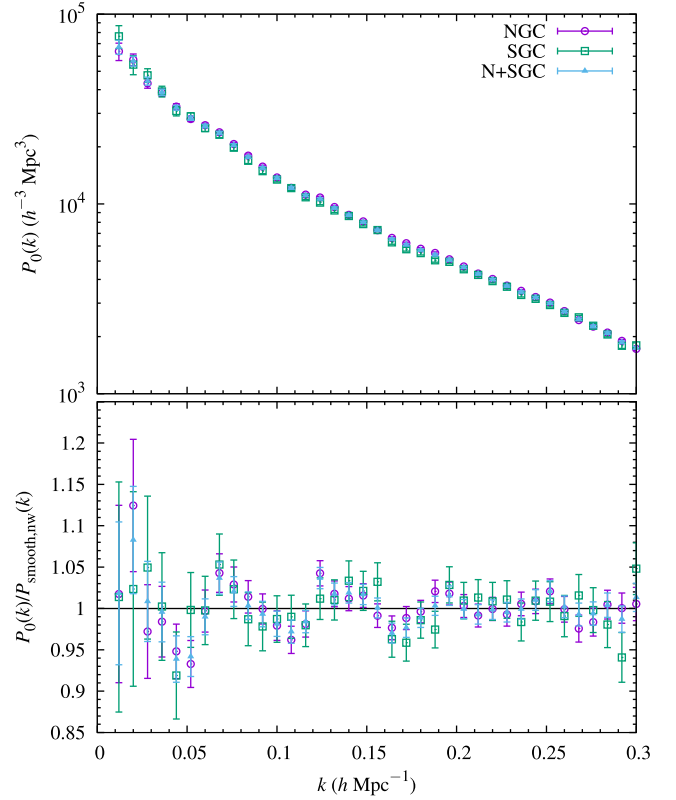


Figure 1. The measured power spectrum from the NGC, SGC, and their volume-weighted average (top panel). To better see how they compare, they are also shown normalized by $P_{\text{smooth},nw}(k)$ as defined in equation (27) in the bottom panel (See the online article for a colour version of this plot.).

2.2 Measuring the galaxy bispectrum

We used the length of the three wave-vectors, k_1 , k_2 , and k_3 , restricted by the triangle condition, $\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = 0$, to parametrize the shape of the bispectrum and averaged over two angles describing the orientation of the triangle (see e.g. Gil-Marín et al. 2015; Gagrani & Samushia 2017). Since we were only using the monopole of the bispectrum we were able to use a Cartesian fast Fourier transform (FFT) without having to worry about the wide-angle effects (Samushia, Branchini & Percival 2015; Scoccimarro 2015).

We employed a ‘brute-force’ algorithm that explicitly looked at all triplets that fell within a specific (k_1, k_2, k_3) bin, of size $0.008h\text{Mpc}^{-1}$ in each k , and satisfied the triangle condition. This algorithm scales as $\mathcal{O}(N_k^2)$ – where N_k is the number of grid points in Fourier space where k is in the range of interest.

It is possible to calculate the bispectrum using point-wise products of FFTs of shells of $\delta(k_1)$, $\delta(k_2)$, and $\delta(k_3)$, with $\delta(k)$ being the FFT of the galaxy overdensity field (see e.g. Scoccimarro 2000; Baldauf et al. 2015; Gil-Marín et al. 2017)

$$\hat{B}(k_1, k_2, k_3) = \frac{V_f}{V_{123}} \int_{\mathbf{x}} \prod_{\kappa=1,2,3} \int_{[q]_{\kappa}} \exp[i\mathbf{q} \cdot \mathbf{x}] \delta(\mathbf{q}), \quad (17)$$

where $V_f = (2\pi)^3/V$, V is the real space volume, and

$$V_{123} \approx \frac{8\pi^2}{(2\pi)^6} k_1 k_2 k_3 \Delta k^3. \quad (18)$$

¹fftw.org

However, we found that it was computationally more efficient to use the ‘brute-force’ algorithm implemented on a graphics processor unit (GPU).²

Our estimate of the bispectrum was (Scoccimarro 2000; Scoccimarro et al. 2001)

$$\hat{B}(k_1, k_2, k_3) = \frac{\langle \Re[\delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\delta(\mathbf{k}_3)] \rangle G(\mathbf{k}_1)G(\mathbf{k}_2)G(\mathbf{k}_3) - S_B}{\int \bar{n}^3(\mathbf{r})w^3(\mathbf{r})d\mathbf{r}} \quad (19)$$

where \Re denotes the real part of a complex number and $\delta(\mathbf{k})$ is the Fourier-transformed overdensity field used in the power spectrum calculation. The shot noise S_B , was computed using the power spectrum estimate of equation (11) (Scoccimarro 2000; Scoccimarro et al. 2001)

$$S_B = (\hat{P}_1 + \hat{P}_2 + \hat{P}_3) \int \bar{n}^2(\mathbf{r})w^3(\mathbf{r})d\mathbf{r} + (1 - \alpha^2) \times \int \bar{n}(\mathbf{r})w^3(\mathbf{r})d\mathbf{r}, \quad (20)$$

where $\hat{P}_i = \hat{P}(k_i)$ is the estimate of the power spectrum at k_i . We computed the combined N+SGC bispectrum in the same way as we computed combined power spectrum, using the weights of equation (15) and simply replacing the power spectrum measurements in equation (16) with the bispectrum measurements.

To implement our estimator on the GPU, we first trimmed the Fourier-transformed overdensity field keeping only the wave numbers in our range of interest, $0.04 \leq k < 0.168$, to better fit in the limited memory resources of the GPUs used for the calculation. We then generated an array of \mathbf{k} values stored as integer multiples of the fundamental frequency in each coordinate direction. Each GPU thread took a single \mathbf{k}_1 and looped over the other \mathbf{k} ’s as \mathbf{k}_2 , starting at \mathbf{k}_1 and going through vectors not yet used as \mathbf{k}_1 , to avoid double counting. The value of \mathbf{k}_3 was then computed from the triangle condition and checked to ensure it was in the range of interest.

For convenience and speed, we used the CUDA types `int4` and `float4` (mocks) or `double4` (data) for storing the \mathbf{k} ’s and the overdensity field, respectively, allowing for four numbers of each type to be stored at a single array index. The `int4` type allowed us to store the three components as well as the corresponding grid index for the overdensity field, which made the lookup times negligible. For the overdensity field, the first two values were the real and imaginary parts, the third value was the magnitude of \mathbf{k} at that grid point, and the fourth value was the correction for the CIC binning associated with that point.

As the GPUs used for this work had relatively low double precision performance – really not much better than a typical CPU – to achieve high throughput while retaining accuracy it was necessary to use mixed precision for calculating the bispectrum from the 2048 NGC and 2048 SGC mocks. Effectively, the calculation of \mathbf{k}_3 was done as integer math, the overall grid correction was calculated using single precision, while the bispectrum contribution was calculated using single precision, but stored and then binned as double precision. The binning was done via a CUDA function `atom-icAdd`. While the double precision version of this function is only supported on NVidia GPUs of compute capability six or higher, the documentation provides an implementation that can be used on GPUs of lower compute capabilities.

²We suspect that a GPU-based implementation of the shell FFT method would be even faster than our $\mathcal{O}(N_k^2)$ method, but could not verify this given the memory limitations (2 GB) of the GPU used for this work.

The reason it is not implemented on those older GPUs, however, is due to its relatively low performance, meaning that it was by far the most expensive step in the process. To reduce the impact of this step, the binning was first done in the GPU thread block shared memory, which has much lower latency than the global GPU memory. Once all the threads in a block went through all of their k_2 ’s, the histogram was then binned into global memory in parallel. Naively, it may seem that the two step process would be less efficient, however, due to the lower latency of the shared memory, reduction in number of writes to global memory, and fewer bank conflicts, you can often see speed up of at least a factor of 2 (Sakharnykh 2015).

Implemented in this manner, we were able to compute the bispectrum from a single mock in ~ 94 s, which was 15 to 20 times faster than our FFT-based implementation, depending on whether a newer Intel core i7 or older AMD FX processor was used, respectively. Our full double precision implementation³ used for processing the data took ~ 310 s. If we had used that implementation to process the mocks, it would have taken approximately 14.5 days to complete. Our mixed-precision implementation cut that down to about 4.5 days.

We show the measured bispectrum from the NGC, SGC, and N+SGC samples in Fig. 2. We also show the measurements normalized by our best-fitting model – see Section 3.2 – to show that it is able to fit the data well.

Unlike the power spectrum, the raw bispectrum measurements do not visibly show the BAO feature. To elucidate the type of signal our constraints come from, we provide a two-dimensional plot of the theoretical bispectrum calculated from a BAO power spectrum normalized by the one calculated from a no-wiggle power spectrum as a function of two wave-vector magnitudes, averaged over the length of the third, in Fig. 3 (see Section 3.2). Plotted in this manner, it is possible to see clear hills and valleys coming from the BAO feature that are equivalent to the decaying oscillatory signature in the power spectrum.

Additionally, we used this plot, along with the need to keep the data vector small enough to ensure reasonable covariance matrices, to decide what bispectrum wave number range to use in our fits. The inset square (cyan in the online version) shows our selected region, which should include the equivalent of the first two ‘wiggles’ in the power spectrum, helping to maximize the BAO constraining power, while keeping the data vector reasonably sized. We note that this two-dimensional plot is merely a convenient way of displaying the BAO features in the bispectrum. Our actual constraints on the scale dilation parameter came from fitting to the three-dimensional data shown in Fig. 2.

2.3 Covariance

We computed the sample covariance from the 2048 MultiDark-PATCHY mocks provided with DR12 (Kitaura et al. 2016; Rodríguez-Torres et al. 2016). This was the main limit to the number of triangles we could use for fitting the bispectrum data. First, we needed to estimate the covariance matrix to enough accuracy that it was not singular, if we were to invert it for our maximum likelihood fitting. Additionally, the errors in our covariance matrix carry through and

³With the recent release of the NVidia Titan V, this time could be reduced to be comparable with the typical time to calculate the power spectrum, e.g. ~ 3 s, for between 1/3 and 1/6 the cost of the comparable, but slightly faster NVidia Tesla V100. This will make studies of the bispectrum substantially quicker in the future.

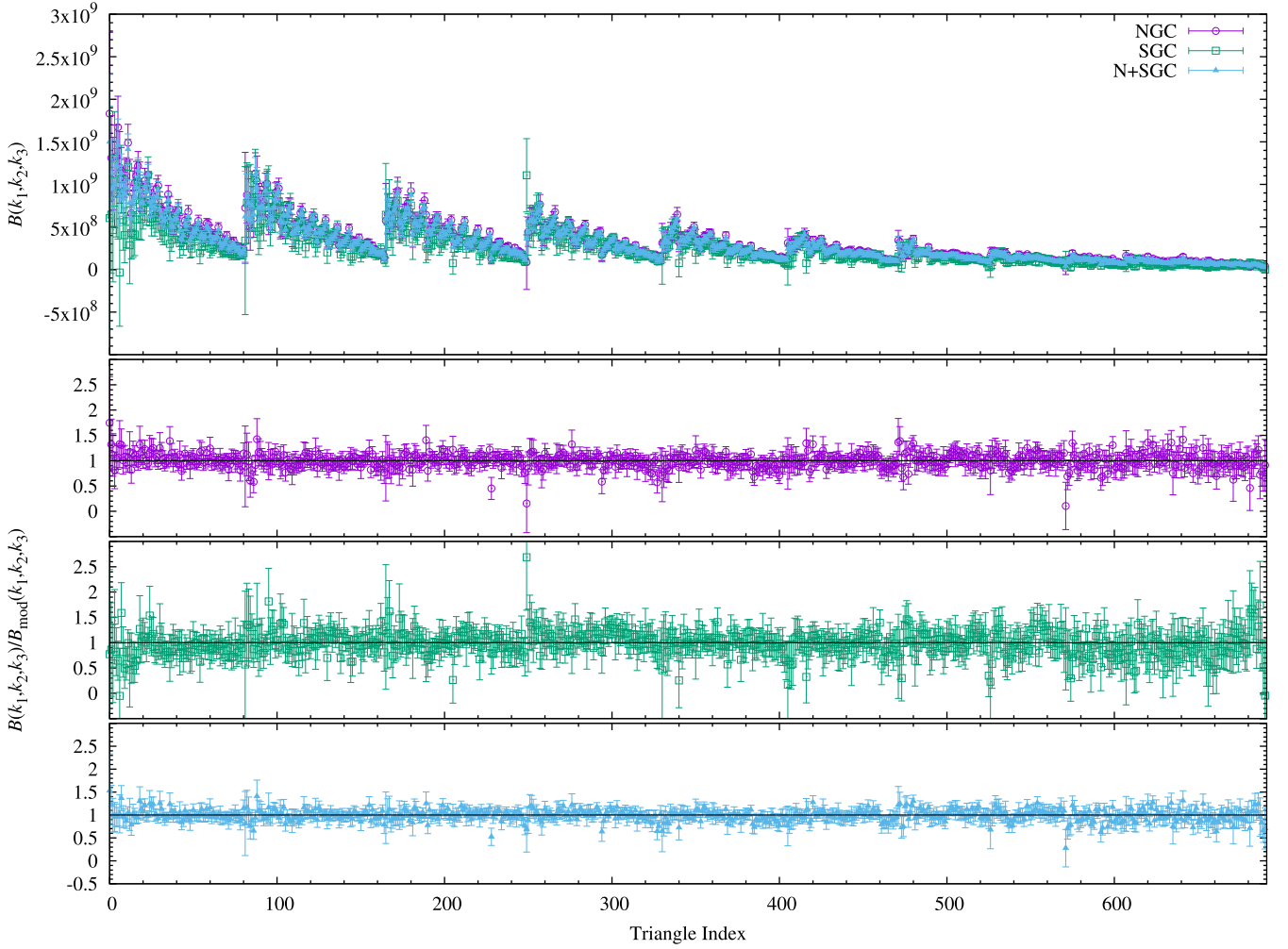


Figure 2. The measured bispectrum from the NGC, SGC, and N+SGC samples. The top panel shows the measurements with uncertainties ($\sqrt{C_{ii}}$) using the same symbols (and colours for the online version) as in Fig. 1. The remaining panels show the data normalized by the best-fitting model. The horizontal axis is the triangle index, which is merely the element number in our data vector. See Appendix A for a more detailed explanation. The fact that the normalized data are simply scattered about one shows that our model accounts for the non-linearities in the data quite well.

affect our constraints on the model parameters. To ensure that the matrix was invertible, and that the uncertainties of its elements were kept low, we limited ourselves to $0.04 \leq k_1, k_2, k_3 \leq 0.168$ for the bispectrum measurements, and $0.008 \leq k < 0.304$ for the power spectrum. In total, we had 691 bispectrum triangles and 37 power spectrum points, which gave at most 728 elements in our data vector.

We ran all the mocks through our data pipelines as described above, then computed the sample covariance,

$$C_{ij} = \frac{1}{n_s - 1} \sum_s (x_i^s - \bar{x}_i)(x_j^s - \bar{x}_j), \quad (21)$$

where n_s is the number of mock samples, and the s in the sum refers to a specific sample. We did this for the NGC, SGC, and N+SGC samples with the power spectrum, bispectrum, and both combined as our data vector, \mathbf{x} . We plot the resulting correlation matrices, $r_{ij} = C_{ij}/(C_{ii}C_{jj})^{1/2}$, in Figs 4 and 5. All seem reasonably well behaved, with more than 90 per cent of the off-diagonal values falling between ± 0.2 .

When using the covariance matrix for our parameter fitting, we corrected for the fact that the inverse of the covariance matrix is a biased estimate of the true inverse covariance needed. This correc-

tion was simply (Hartlap, Simon & Schneider 2007; Percival et al. 2014)

$$\Psi = \left(\frac{n_s - n_b - 2}{n_s - 1} \right) \mathbf{C}^{-1} \quad (22)$$

where n_b is the number of values in our data vector. Given that $n_s = 2048$, and $n_b = 37, 691$, and 728 for the power spectrum, bispectrum, and combined data vectors, respectively, the above correction factor was then 0.981, 0.662, and 0.644, which means our constraints on parameters from the bispectrum and the combined data were affected by the limited number of mock catalogues.

In addition to the correction above, we also had to carry through the uncertainty in the covariance matrix elements themselves due to limited number of mock catalogues (Taylor, Joachimi & Kitching 2013). This affected the variance of the parameters being constrained, increasing them by a factor of (Percival et al. 2014)

$$m_1 = \frac{1 + \beta(n_b - n_p)}{1 + A + \beta(n_p + 1)}, \quad (23)$$

where

$$A = \frac{2}{(n_s - n_b - 1)(n_s - n_b - 4)}, \quad (24)$$

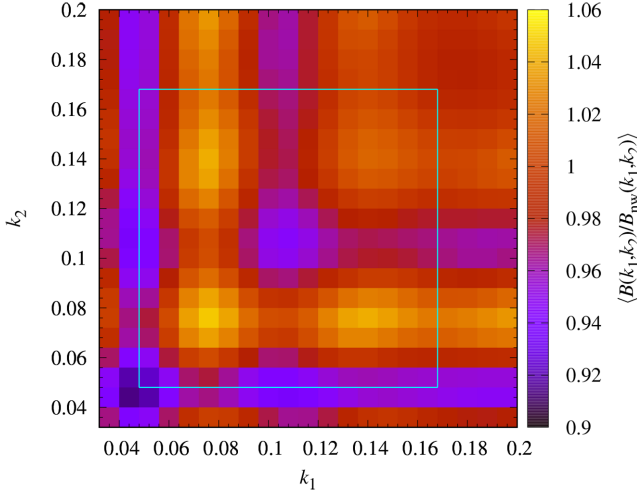


Figure 3. The theoretical bispectrum normalized by the theoretical no-wiggle bispectrum, as a function of the length of two wave-vectors, averaged over the length of the third wave-vector. The inset box (cyan in the online version) encloses the scales used in our analysis. The two-dimensional sequences of local maxima and minima are manifestations of the BAO feature. (See the online article for a colour version of this plot.)

$$\beta = \frac{(n_s - n_b - 2)}{(n_s - n_b - 1)(n_s - n_b - 4)}, \quad (25)$$

and n_p is the number of parameters in the model being fit to the data. For our analysis we found values of $m_1 = 1.0089, 1.4982$, and 1.5248 for the power spectrum, bispectrum, and combined sample, respectively. This substantially reduced the potential constraints from the bispectrum and the combined samples, highlighting a need for either a larger number of mock catalogues or a high-precision theoretical model of the covariance. We note that the assumptions under which the above corrections were derived may not hold for the bispectrum, and should be explored in more detail in future works.

3 MODEL AND FITTING

3.1 The power spectrum

To model the power spectrum, we followed the method of Anderson et al. (2014), with one modification. We used the Python implementation of CAMB (Lewis, Challinor & Lasenby 2000) to generate a linear power spectrum with our fiducial cosmology. Then, using the fitting formula of Eisenstein & Hu (1998), we calculated a no-wiggle power spectrum to match the broad-band shape of the linear power spectrum from CAMB. These are convolved with the window function of the survey before their use in the model below. Since the data contained non-linearities, we used a smoothing polynomial to match the linear theory power spectrum to the data,

$$\mathcal{P}(k) = a_0 k^2 + a_1 k + a_2 + \frac{a_3}{k} + \frac{a_4}{k^2} + \frac{a_5}{k^3}. \quad (26)$$

This differs from the method used in Anderson et al. (2014) by including a term proportional to k^2 . We found that including this extra term gives us a better fit.⁴ We added this polynomial to the

no-wiggle power spectrum

$$P_{\text{smooth,nw}}(k) = B^2 P_{\text{nw}}(k) + \mathcal{P}(k), \quad (27)$$

where B is an amplitude parameter to account for galaxy bias and gravitational growth. This was then multiplied by an oscillatory part

$$O(k) = \left[1 + \left(\frac{P_{\text{lin}}(k/\alpha)}{P_{\text{nw}}(k/\alpha)} - 1 \right) \exp \left(\frac{1}{2} \Sigma^2 k^2 \right) \right] \quad (28)$$

where α is defined in equation (1), and Σ is a Finger-of-God damping parameter. This gave our final power spectrum model as

$$P_{\text{mod}}(k) = P_{\text{smooth,nw}}(k) O(k). \quad (29)$$

Fig. 6 shows the average power spectrum of the 2048 mocks divided by the no-wiggle best-fitting model – e.g. the average divided by equation (27) with the best-fitting values of the parameters – as well as equation (28) with the best-fitting parameter values.

For the fitting, this model had nine free parameters, the a_i of equation (26), the amplitude parameter, B , the non-linear damping parameter, Σ , and the scale dilation parameter, α . We used a custom Markov Chain Monte Carlo (MCMC) code that utilized the Metropolis–Hastings algorithm (Hastings 1970). This choice was made so that when fitting simultaneously to the power spectrum and bispectrum, the exact same code and algorithm was used for both, as the bispectrum fitting had to be done with custom code in order to utilize a GPU to speed up the model calculation (see Section 3.2).

3.2 The bispectrum

To model the galaxy bispectrum we used the second-order perturbation theory model presented by Scoccimarro (2000), with some small changes to account for non-linearities in the data. The first- and second-order kernels are

$$Z_1(\mathbf{k}) = (b_1 + f\mu^2) \quad (30)$$

and

$$Z_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{b_2}{2} + b_1 F_2(\mathbf{k}_1, \mathbf{k}_2) + f\mu^2 G_2(\mathbf{k}_1, \mathbf{k}_2) + \frac{f\mu k}{2} \left[\frac{\mu_1}{k_1} Z_1(\mathbf{k}_2) + \frac{\mu_2}{k_2} Z_1(\mathbf{k}_1) \right], \quad (31)$$

where $\mu \equiv \mathbf{k} \cdot \hat{\mathbf{z}}/k$, with $\mathbf{k} \equiv \mathbf{k}_1 + \mathbf{k}_2$, $\mu_i \equiv \mathbf{k}_i \cdot \hat{\mathbf{z}}/k_i$,

$$F_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{5}{7} + \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{2k_1 k_2} \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right) + \frac{2}{7} \left(\frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2} \right)^2, \quad (32)$$

and

$$G_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{3}{7} + \frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{2k_1 k_2} \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right) + \frac{4}{7} \left(\frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2} \right)^2. \quad (33)$$

From these, our model bispectrum was

$$B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = [2Z_2(\mathbf{k}_1, \mathbf{k}_2)Z_1(\mathbf{k}_1)Z_1(\mathbf{k}_2)P(k_1)P(k_2) + \text{cyc.}] D_{\text{FoG}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3), \quad (34)$$

where

$$D_{\text{FoG}}(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = \frac{1}{1 + (k_1^2 \mu_1^2 + k_2^2 \mu_2^2 + k_3^2 \mu_3^2) \sigma_v^2 / 2} \quad (35)$$

and $P(\mathbf{k})$ is the non-linear power spectrum calculated for our fiducial cosmology from CAMB that was also convolved with the survey window function before use in the above model. This treatment of the window function for the bispectrum model is consistent with that of Gil-Marín et al. (2015), which is an adequate approximation so long as all three sides are sufficiently larger than the fundamental

⁴This was due to us accidentally multiplying the shot noise by the CIC grid correction. Thus, including this term will have negated that mistake.

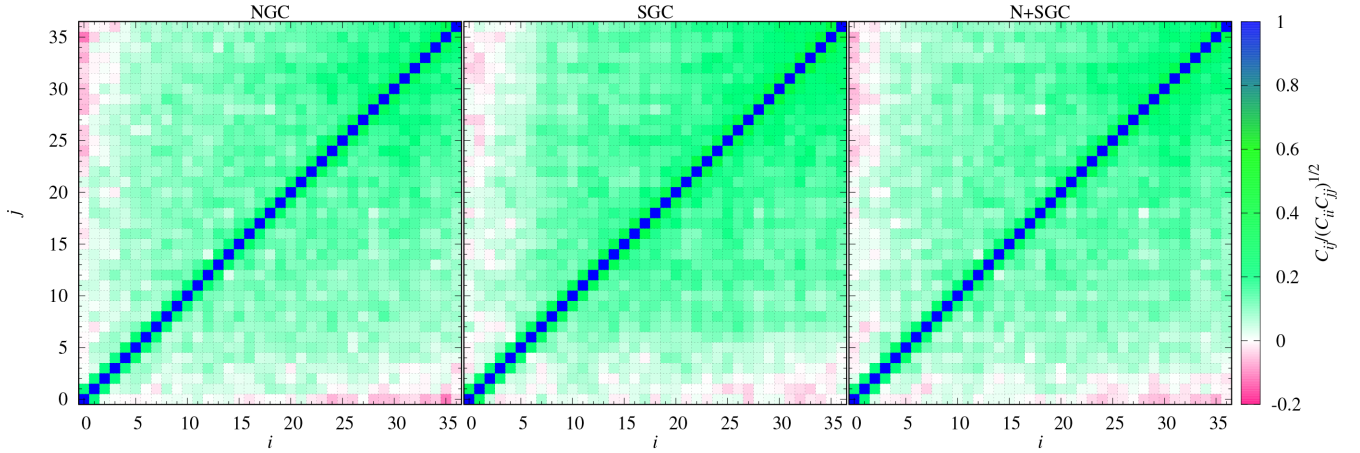


Figure 4. The correlation matrices for the power spectrum measurements calculated from 2048 Patchy mocks. The power spectra were calculated from the NGC and SGC, then a weighted average. Most of the off-diagonal values vary between ± 0.2 . (See the online article for a colour version of this plot.)

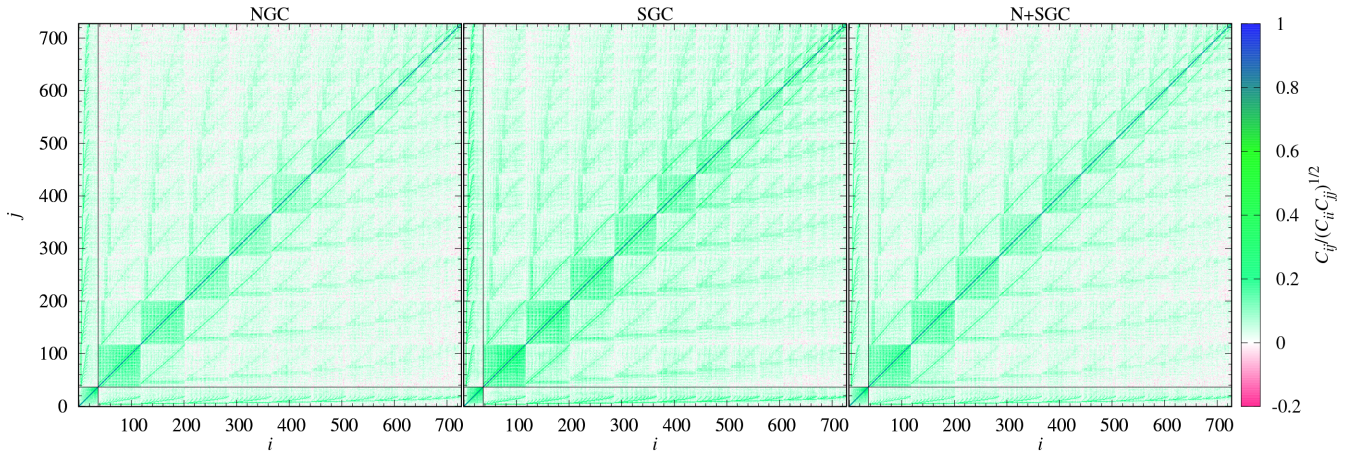


Figure 5. The correlation matrices for the power spectrum and bispectrum with cross-correlations calculated from the 2048 Patchy mocks. The vertical and horizontal solid lines divide the matrix into the power spectrum (bottom left, small square), the bispectrum (top right, large square), and their cross-correlations (vertical rectangle on the left, or horizontal rectangle on the bottom). The off-diagonal elements still mostly only vary between ± 0.2 . (See the online article for colour a version of this plot.)

frequency, which is $\sim 0.005 \, h \text{ Mpc}^{-1}$ for the NGC and $\sim 0.007 \, h \text{ Mpc}^{-1}$ for the SGC.

The addition of the Finger-of-God suppression factor and the use of a non-linear power spectrum were the changes we made to better fit the data. Even though this model has been shown to be inadequate for fitting the full non-linear bispectrum shape (Saito et al. 2014; Gil-Marín et al. 2015, 2017), we found it to be adequate for the range of k that we consider (e.g. $0.04 \leq k \leq 0.168$) in our analysis. Since we were only interested in the position of the BAO peak we could afford to marginalize over smooth systematic effects in the bispectrum shape without properly modelling them. We found that this was enough to get unbiased estimates of α and no extra smoothing polynomials seemed to be necessary to achieve a good fit to the bispectrum data (see the lower panels of Fig. 2).

Since we were examining the bispectrum monopole, equation (34) was spherically averaged by integrating over two angles: the angle of \mathbf{k}_1 with the line of sight, μ_1 , and the azimuthal angle of \mathbf{k}_2 around \mathbf{k}_1 , φ (see e.g. Gagrani & Samushia 2017, section 3.1). This

gave the model as

$$B(k_1, k_2, k_3) = \frac{1}{4\pi} \int_{-1}^1 d\mu_1 \int_0^{2\pi} d\phi B(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3). \quad (36)$$

The values of μ_2 and μ_3 could easily be calculated from the value of the two angles above. We took the Alcock–Paczynski effect (Alcock & Paczynski 1979; Kaiser 1987; Ballinger, Peacock & Heavens 1996; Simpson & Peacock 2010; Samushia et al. 2011) into account by transforming the measured k 's and μ 's as

$$k_i \rightarrow \frac{k_i}{\alpha_{\perp}} \left[1 + \mu_i^2 \left(\frac{\alpha_{\perp}^2}{\alpha_{\parallel}^2} - 1 \right) \right]^{1/2}, \quad (37)$$

$$\mu_i \rightarrow \frac{\alpha_{\perp}}{\alpha_{\parallel}} \mu_i \left[1 + \mu_i^2 \left(\frac{\alpha_{\perp}^2}{\alpha_{\parallel}^2} - 1 \right) \right]^{-1/2}, \quad (38)$$

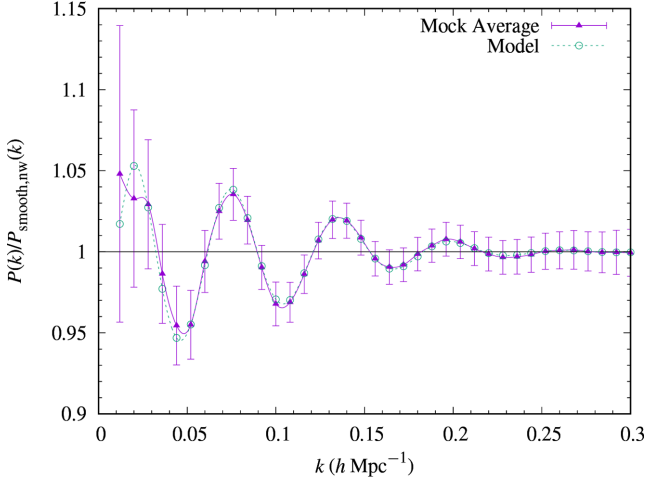


Figure 6. A comparison of the best-fitting model and the average-measured power spectrum from mock galaxy catalogues. The triangles show the mock average normalized by $P_{\text{smooth},nw}(k)$ defined in equation (27), calculated with the best-fitting values of the a_i , and B parameters, with a smooth cubic spline drawn through the points. The open circles show the model of equation (29) with the best-fitting parameters normalized again by $P_{\text{smooth},nw}(k)$. The two agree quite well, particularly with respect to the BAO peak positions.

where

$$\alpha_{\parallel} = \frac{H^{\text{fid}}(z)r_d^{\text{fid}}}{H(z)r_d}, \quad (39)$$

and

$$\alpha_{\perp} = \frac{D_A(z)r_d^{\text{fid}}}{D_A^{\text{fid}}(z)r_d}, \quad (40)$$

along with renormalizing the power spectrum by a factor of $1/\alpha_{\perp}^2\alpha_{\parallel}$ and due to equation (34), the bispectrum by the same factor squared. From equations (1), (3), (39), and (40), it can be seen that α_{\parallel} and α_{\perp} are related to α via

$$\alpha^3 = \alpha_{\perp}^2\alpha_{\parallel}. \quad (41)$$

Since the double integral of equation (36) had to be evaluated for each of our 691 k -triplets a very large number of times for the MCMC fitting procedure, it had to be implemented in a numerically efficient manner. For this we again turned to the GPU allowing us to calculate the double integral using a couple of levels of parallelism.

While it is possible to implement adaptive quadrature on the GPU, the error estimation steps can introduce significant overhead, and most algorithms rely on recursion that is not well suited for GPUs (Thuerck et al. 2014). Given this, we instead opted for the much easier to implement, fixed Gaussian quadrature rules. Since Gaussian quadrature can give an exact result for polynomials of degree $2n - 1$ or less, it can allow very accurate numerical integration with relatively few function evaluations. Given that the exact shape of the above integrand can be difficult to predict, keeping n as large as possible was desirable. Additionally, we again ran into the fact that commodity GPUs achieve the highest throughput for single precision floating-point calculations. This made the use of mixed precision necessary, where many of the calculations are done and variables stored as single precision floats.

Extending Gaussian quadrature in two dimensions was done simply by setting up a two-dimensional grid such that

$$\int f(x, y) dx dy \cong \sum_{i=1}^n \sum_{j=1}^n w_i w_j f(x_i, y_j), \quad (42)$$

where x_i are the points to evaluate your function determined from your Gaussian quadrature rule, and w_i are their associated weights, both of which can be readily found in handbooks. Combine the need for a two-dimensional grid, the unknown shape of the integrand, the need to use mixed precision and the fact that the maximum number of threads per GPU thread block is 1024, and $n = 32$ becomes a natural choice.

This allowed us to have one thread block per k -triplet, where the integral was then approximated by a two-dimensional (32×32) -point Gaussian quadrature rule. Each thread then computed one contribution to the integral, and stored the result in block-shared memory, with the final summing done in a two-step reduction. To reduce the impact of mixed precision, we stored all the calculations of equations (30) – (33) as single precision and the calculations of equation (34) and (35) as double precision. We then perform the final summing over the two-dimensional grid using those double precision values and return the result as double precision. In our tests, a complete double precision calculation using the exact same algorithm has a relative difference – e.g. $(B_{\text{MP}} - B_{\text{DP}})/B_{\text{DP}}$ – from our mixed-precision calculation of $\sim 10^{-7}$. Given the relatively large uncertainties in the measured bispectrum, this loss of precision was well worth ~ 12 times speed-up of the model evaluation.

For the fitting, we used six free parameters: the three from equations (30) and (31), e.g. b_1 , linear bias, b_2 , second-order bias, and f , the linear growth factor⁵ with the two Alcock–Paczynski effect (Alcock & Paczynski 1979; Kaiser 1987; Ballinger et al. 1996; Simpson & Peacock 2010; Samushia et al. 2011) parameters, α_{\parallel} and α_{\perp} , and the Finger-of-God velocity dispersion parameter, σ_v^2 . Since our model was only validated for the purposes of measuring the BAO feature we did not attach any cosmologically meaningful interpretation to the estimates of the parameters b_1 , b_2 , f , or σ_v^2 . They were reasonably close to the linear model expectations but were very likely strongly affected by systematics and we therefore do not quote them as useful cosmological constraints in this work.

Since we were only fitting to the spherically averaged bispectrum monopole, we didn’t expect to be able to reliably constrain both α_{\parallel} and α_{\perp} . Instead, as in the case of the spherically averaged two-point statistics (see e.g. Anderson et al. 2012, 2014), we expected to only constrain the value of the single scale dilation parameter, α . We elected to keep the fit in terms of α_{\parallel} and α_{\perp} . Then we calculated the value of α via equation (41) for each of our accepted parameter realizations in the MCMC chain. It is not immediately obvious that the best-constrained combination of α ’s in the bispectrum is the same as for the power spectrum. By inspecting our MCMC chains we were able to verify that the two are in fact very close. Fig. 7 shows one of our MCMC chains projected onto the α_{\parallel} – α_{\perp} plane and the principle axis of the likelihood ellipse is very well aligned with the direction of α in equation (41).

We ran our MCMC chains with no priors aside from the loose requirement the f remain positive. We used our own custom MCMC software utilizing the Metropolis–Hastings algorithm (Hastings 1970), which was easier to interface with the model calculation

⁵These parameters can only be measured up to some overall power spectrum normalization, σ_8 , which we leave off for brevity. In the text, when we use b_1 , b_2 , or f , we mean the combinations $b_1\sigma_8$, $b_2\sigma_8$, or $f\sigma_8$

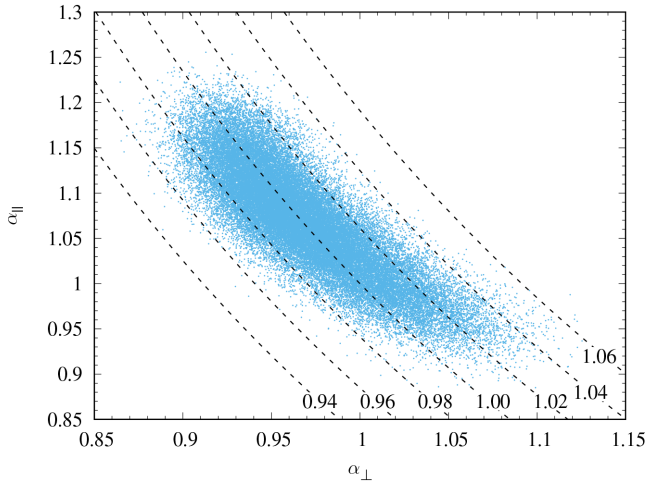


Figure 7. Accepted realizations from one of our MCMC chains fitting to the mock average N+SGC bispectrum projected onto α_{\parallel} - α_{\perp} plane. The dashed lines correspond to fixed values of $\alpha^3 \equiv \alpha_{\perp}^2 \alpha_{\parallel}$.

on the GPU. Given that only one GPU was available for the model calculation and the GPU's constant memory was used to store the values of the parameters for each model evaluation, we simply ran a single chain for many millions of realizations instead of running multiple, simultaneous chains needed to test for convergence. We built our code with the option to resume a chain should our post-processing reveal that it had not sufficiently explored the parameter space.

3.3 The combined power spectrum and bispectrum model

For the joint fit our data vector simply became the 37 power spectrum values, followed by the 691 bispectrum values. The model was then simply calculating the first 37 values using the power spectrum model of Section 3.1 and the next 691 values with the bispectrum model of Section 3.2.

The difference came in the free parameters, particularly α , α_{\perp} , and α_{\parallel} . We only let α_{\perp} and α_{\parallel} vary freely with the α for the power spectrum model then being fixed by equation (41). In all, for the combined model we ended up with 14 free parameters: B , Σ , $a_0 - a_5$, b_1 , b_2 , f , α_{\perp} , α_{\parallel} , and σ_v . This likely could have been reduced by relating the power spectrum amplitude parameter B to the biases, b_1 and b_2 , and the linear growth factor, f . However, since we were only concerned with the constraints on α in the end, all other parameters were treated as nuisance parameters, and marginalized against anyway. We again ran our MCMC chains with only loose constraints to ensure that parameters that enter the models as squares remained positive.

4 RESULTS

We performed a number of fittings to the power spectrum, bispectrum, and their combination, analysing the NGC and SGC samples separately before fitting to the volume-weighted averages. This was done to ensure that the results from any one fitting were consistent with the results from others. We additionally fit the models to the average power spectrum and bispectrum from the mock catalogues for the purposes of verifying our analysis pipeline and assessing systematic errors.

Table 1. The results of the MCMC fittings for the BAO scale parameter, α . Both the values measured from the mocks and the data for the two galactic cap samples, as well as the volume-weighted averages, are presented. The standard deviations have been scaled due to the covariance uncertainty.

Data	Sample	α (mock)	α (DR12)
$P(k)$	NGC	1.007 ± 0.018	0.976 ± 0.012
	SGC	1.003 ± 0.025	1.025 ± 0.021
	N+SGC	1.003 ± 0.013	0.988 ± 0.012
$B(k_1, k_2, k_3)$	NGC	1.002 ± 0.020	0.978 ± 0.031
	SGC	1.007 ± 0.030	1.086 ± 0.072
	N+SGC	1.001 ± 0.017	0.991 ± 0.027
$P(k) + B(k_1, k_2, k_3)$	NGC	1.001 ± 0.016	0.982 ± 0.011
	SGC	1.005 ± 0.020	1.020 ± 0.020
	N+SGC	1.002 ± 0.010	0.988 ± 0.011

Our fiducial cosmology was virtually identical to that of the MultiDark-PATCHY mock catalogues. As such, we would expect that α should have been very nearly one when fit to the ensemble of mocks – the exact expected value is $\alpha_{\text{mock}} = 1.00005$. Looking at Table 1, we can see that this indeed turned out to be the case. We take the largest per cent deviation from the expected value of α for each of the fittings – e.g. the power spectrum, bispectrum, and joint fittings – to indicate our systematic error. These errors were ~ 0.7 per cent for the power spectrum and bispectrum, and ~ 0.5 per cent for the combined fitting.

For the DR12 measurements, our results for the NGC and SGC separately are in agreement with Ross et al. (2017), who also analysed the NGC and SGC separately, finding a slightly lower value for the NGC and a slightly larger one for the SGC, and the combined result being closer to one. We also note the agreement of our power spectrum and bispectrum results. The largest difference is for the SGC measurements, which likely had to do with the smaller volume resulting in a noisier measurement of the bispectrum.

We show the histograms for α from all of our MCMC fittings in Figs 8, 9, and 10 for the N+SGC, NGC, and SGC samples, respectively. In all the figures, the results of fitting to the mocks is shown on the left, and the fitting to the data on the right, and all histograms have been normalized so that the area under the curves is equal to one. All the histograms are very close to Gaussian and encapsulate all regions of relatively high likelihood suggesting that α parameter space was well explored by our MCMC chains. These plots have not been broadened by the m_1 factor of equation (23).

The standard deviations listed in Table 1 have been broadened by the m_1 factors. The effect is quite apparent in the full sample, N+SGC, combined power spectrum-bispectrum fitting. Without carrying the covariance uncertainty through, the standard deviation would be 0.0087, a ~ 27 per cent tighter constraint than the power spectrum alone, and a better constraint than the one from the power spectrum of the reconstructed field. However, after multiplying by the square root of m_1 , this become 0.0108 that still represents a ~ 10 per cent tighter constraint than our power spectrum fitting.

To test how the constraints on α may improve given a more precise estimation of the covariance matrix, we ran chains fitting to the average of the mocks with the assumption that our covariance was drawn from one million mock catalogues, which would make the effects of both the covariance scaling and m_1 factor negligible. We did this for both the bispectrum and joint fitting, finding that the standard deviation for α dropped to 0.011 and 0.007 from 0.014 and 0.010, respectively. This indicates that given a less noisy estimate of the covariance, from either more mock catalogues, a high-precision theoretical calculation, or a hybrid approach like the one used by

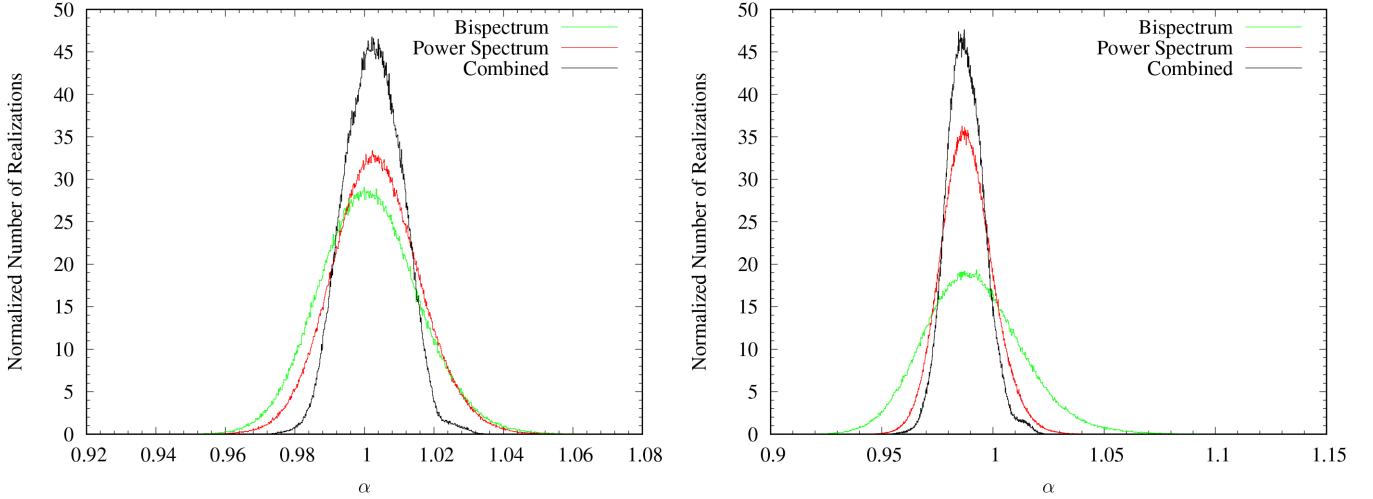


Figure 8. Histograms of the MCMC realizations from fitting to the bispectrum only, the power spectrum only, and the combination for the N+SGC. The left-hand panel shows the results from fitting to the average of the mocks. The right-hand panel shows the results from fitting to the measurements from the data. (See the online article for a colour version of this plot.)

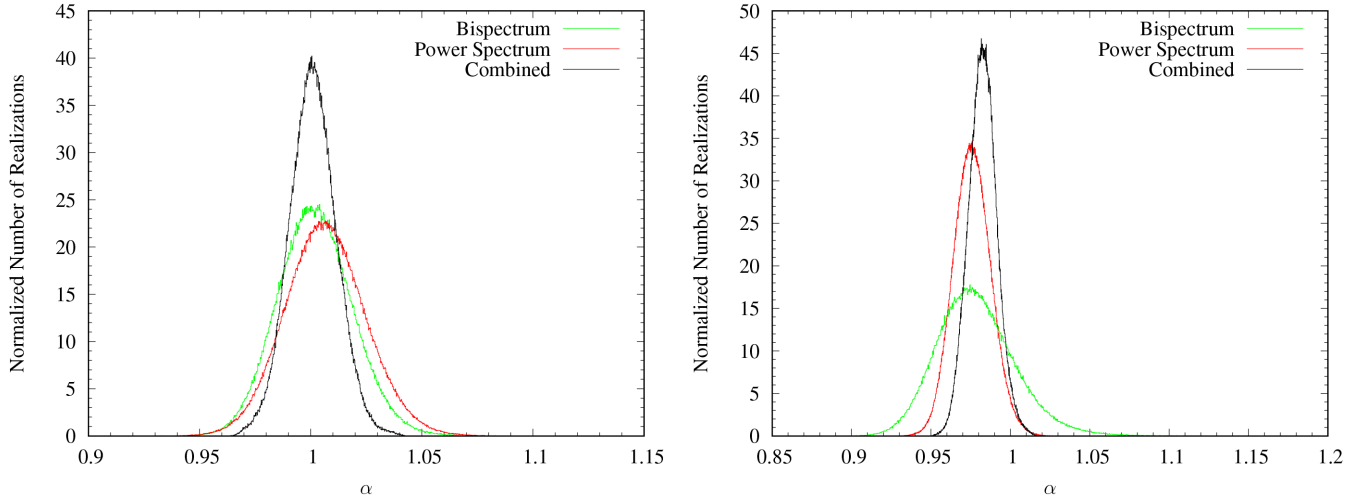


Figure 9. Histograms of the MCMC realizations from fitting to the bispectrum only, the power spectrum only, and the combination for the NGC CMASS sample. The panels are the same as in Fig. 8. (See the online article for a colour version of this plot.)

Slepian et al. (2017a,b), could improve constraints on the distance scale by 30 per cent.

We note that we achieve a significantly better fit by using a ‘wiggly’ power spectrum in our bispectrum model than a ‘no-wiggly’ power spectrum, with a χ^2 penalty for the no-wiggly model of $\Delta\chi^2 = 20.64$. This implies a 4.1σ detection of the BAO features in the BOSS DR12 galaxy bispectrum, comparable to the significance of the detection by Slepian et al. (2017b).

Of course, α itself is merely a means to measure the distance to the survey redshift via equation (2). First it was necessary to compute $D_V^{\text{fid}}(z)$, which requires calculating $D_A^{\text{fid}}(z)$, and $H^{\text{fid}}(z)$. For our fiducial cosmology, these come out to $H^{\text{fid}}(z = 0.57) = 93.04 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $D_A^{\text{fid}}(z = 0.57) = 1386.01 \text{ Mpc}$, and $D_V^{\text{fid}}(z = 0.57) = 2056.45 \text{ Mpc}$. We also note that the drag radius for our fiducial cosmology was $r_d^{\text{fid}} = 147.59 \text{ Mpc}$ as calculated from the fitting formula of Hu & Sugiyama (1996) and Eisenstein & Hu (1998). The values of D_V for the various fittings are reported in Table 2, where the r_d^{fid}/r_d was omitted from the column headings for brevity.

We take the measurement from the combined power spectrum plus bispectrum fitting to the N+SGC data as our main result,

$$D_V(z = 0.57) = 2031 \pm 22 \text{ Mpc} \left(\frac{r_d}{r_d^{\text{fid}}} \right). \quad (43)$$

We compared our result to that of other works to test its robustness, after converting those results to be consistent with our fiducial cosmology. Anderson et al. (2014) found

$$D_V^{\text{A14}}(z = 0.57) = 2032 \pm 20 \text{ Mpc} \left(\frac{r_d}{r_d^{\text{fid}}} \right), \quad (44)$$

from their analysis of the DR11 CMASS sample. In analysing the DR12 CMASS sample Cuesta et al. (2016) found

$$D_V^{\text{C16}}(z = 0.57) = 2035 \pm 20 \text{ Mpc} \left(\frac{r_d}{r_d^{\text{fid}}} \right), \quad (45)$$

Gil-Marín et al. (2016) found

$$D_V^{\text{G16}}(z = 0.57) = 2035 \pm 19 \text{ Mpc} \left(\frac{r_d}{r_d^{\text{fid}}} \right), \quad (46)$$

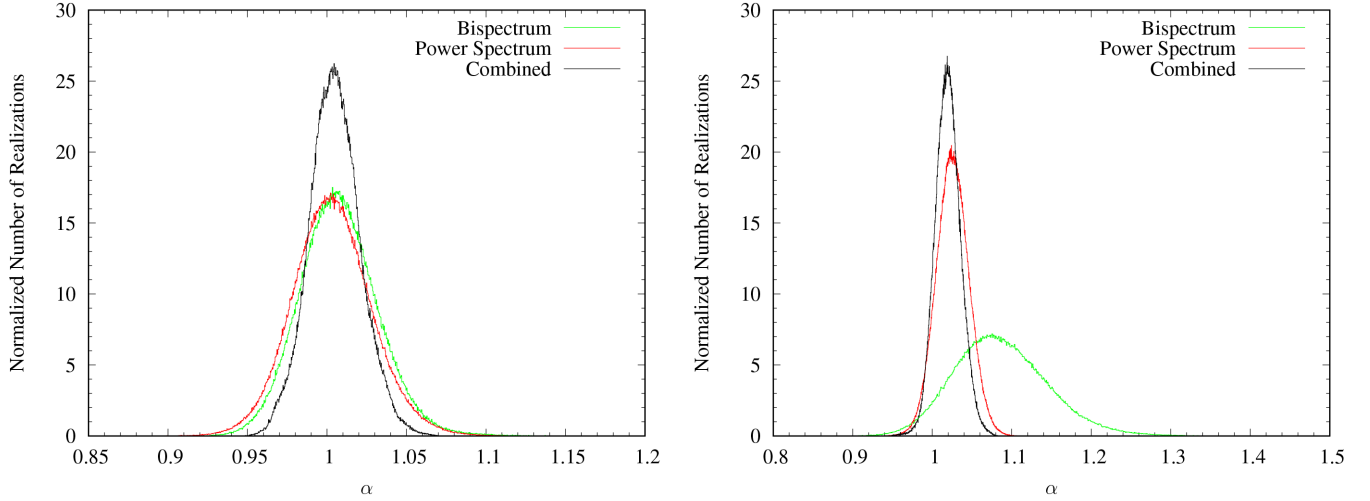


Figure 10. Histograms of the MCMC realizations from fitting to the bispectrum only, the power spectrum only, and the combination for the SGC CMASS sample. The panels are the same as in Fig. 8. (See online article for colour version of this plot.)

Table 2. The D_V values from the various fittings in Table 1. While the columns are labeled as D_V we note that they actually represent $(r_d^{\text{fid}}/r_d)D_V$ values. The values for DR12 reported here agree remarkably well with the values reported by other analyses, and the values from the fittings to the mocks agree with our expectations.

Data	Sample	D_V (mock) (Mpc)	D_V (DR12) (Mpc)
$P(k)$	NGC	2070 ± 38	2007 ± 24
	SGC	2063 ± 51	2108 ± 42
	N+SGC	2063 ± 26	2032 ± 24
$B(k_1, k_2, k_3)$	NGC	2060 ± 42	2012 ± 63
	SGC	2071 ± 62	2233 ± 148
	N+SGC	2059 ± 36	2038 ± 55
$P(k) + B(k_1, k_2, k_3)$	NGC	2058 ± 27	2020 ± 23
	SGC	2066 ± 41	2097 ± 41
	N+SGC	2061 ± 21	2031 ± 22

Ross et al. (2017) found,

$$D_V^{\text{R17}}(z = 0.57) = 2022 \pm 17 \text{ Mpc} \left(\frac{r_d}{r_d^{\text{fid}}} \right), \quad (47)$$

and Slepian et al. (2017b) found

$$D_V^{\text{S17m}}(z = 0.57) = 2036 \pm 33 \text{ Mpc} \left(\frac{r_d}{r_d^{\text{fid}}} \right), \quad (48)$$

from their ‘minimal’ model and

$$D_V^{\text{S17t}}(z = 0.57) = 2026 \pm 29 \text{ Mpc} \left(\frac{r_d}{r_d^{\text{fid}}} \right), \quad (49)$$

from their ‘tidal’ model. Our main result deviates the most from that of Ross et al. (2017), and even then the disagreement is within $\sim 0.5\sigma$.

Additionally, Ross et al. (2017) gave results for the NGC and SGC separately, allowing for a more detailed comparison. Their fitting to the two-point correlation function gave

$$D_V^{\text{R17,NGC}}(z = 0.57) = 2006 \pm 21 \text{ Mpc} \left(\frac{r_d}{r_d^{\text{fid}}} \right), \quad (50)$$

and

$$D_V^{\text{R17,SGC}}(z = 0.57) = 2090 \pm 41 \text{ Mpc} \left(\frac{r_d}{r_d^{\text{fid}}} \right). \quad (51)$$

These values agree remarkably well with our analysis, showing a somewhat lower value from the NGC and a higher value from the SGC.

Similar to Slepian et al. (2017b), we assessed our systematic errors by examining the bias of our results of fitting to the mocks. We take the largest deviation from the expected value of α in each of the fittings, finding systematic errors of ± 15 Mpc for the power spectrum and bispectrum fittings, and ± 10 Mpc for the combined fitting. This gives our main result as

$$D_V(z = 0.57) = 2031 \pm 22 \text{ (stat.)} \pm 10 \text{ (sys.) Mpc} \left(\frac{r_d}{r_d^{\text{fid}}} \right). \quad (52)$$

5 CONCLUSIONS

We report an independent measurement of the distance to the BOSS DR12 CMASS sample of $D_V(z = 0.57) = 2032 \pm 24 \text{ Mpc}(r_d/r_d^{\text{fid}})$ from the power spectrum, $D_V(z = 0.57) = 2038 \pm 55 \text{ Mpc}(r_d/r_d^{\text{fid}})$ from the bispectrum, and $D_V(z = 0.57) = 2031 \pm 22 \text{ Mpc}(r_d/r_d^{\text{fid}})$ from the combined analysis. These values are in remarkable agreement with each other, and with the analyses of Anderson et al. (2014), Cuesta et al. (2016), Gil-Marín et al. (2016), Ross et al. (2017), and Slepian et al. (2017b). The power spectrum gives a ~ 1.2 per cent constraint (~ 1.4 per cent with systematics), and the bispectrum gives a ~ 2.7 per cent constraint (~ 2.8 per cent with systematics). The combined analyses gives a ~ 1.1 per cent constraint (~ 1.2 per cent with systematics), mainly limited by the number mocks available for covariance estimation. However, when combined the constraint still improves by ~ 10 per cent compared to the power spectrum only constraints.

Our bispectrum constraints from the mocks were tighter than the ones from the data, suggesting that either this specific realisation of the DR12 CMASS volume is slightly noisier than typical from the point of view of the bispectrum monopole estimator, or that the mocks themselves do not fully capture the cosmological variance of the bispectrum. When fitting to the mean of the mocks we get a 1.7 per cent constraint from the bispectrum only and a 1 per cent

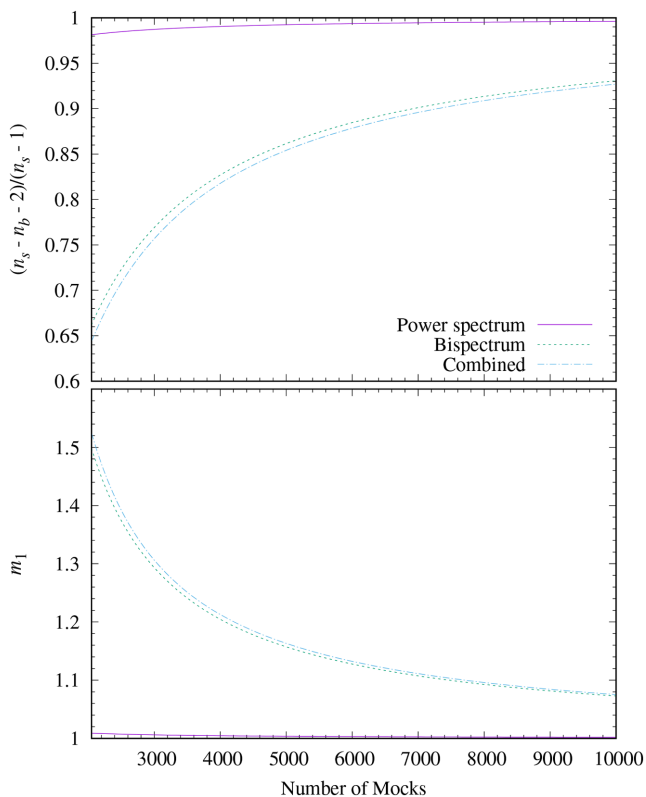


Figure 11. The effects of the limited number of mock catalogues. The top panel shows the factor that multiplies the inverse sample covariance, and the bottom panel shows how much the variance of parameter constraints should be inflated. The solid line (purple online) shows the quantities for the power spectrum, the dashed line (green online) shows them for the bispectrum, and the dot-dashed line (cyan online) shows them for the combined data vector. The horizontal axis starts at the number of mocks used in this work. It is clear that increasing the number of mocks will have little effect on the power spectrum, but a dramatic effect on bispectrum and combined data.

constraint from the joint fit to the power spectrum and bispectrum. The numbers in Table 1 suggest that if a better model for the covariance were available the joint constraints from the power spectrum and the bispectrum would be comparable to the constraints from reconstructed power spectrum even at the current level of systematics.

The main limiting factor to the precision of the bispectrum measurements is a relatively small number of mock catalogues available for the covariance estimation. Having only 2048 mocks means that the values in our inverse covariance matrix estimate were all substantially reduced – by a factor of 0.662 for the bispectrum and 0.644 for the combined data – leading to broader posterior likelihoods. Having a better estimate of the bispectrum and joint covariance matrices would reduce the error on the α parameter from the joint fit by an extra ~ 30 per cent. These improved estimates could come from either a larger number of mocks or analytic calculations.

Looking at Fig. 11, we can clearly see that the power spectrum constraints will not really benefit from an increased number of mock catalogues, while the constraints from the bispectrum and the combined data can see dramatic improvements with $\sim 10\,000$ mocks. With $\sim 40\,000$ mocks, the additional uncertainty due to noise in the inverse covariance matrix would become negligible.

However, generating that number of mock catalogues is a computationally expensive proposition, which is only going to be ex-

acerbated by the increased volumes and number densities of future redshift surveys such as, the upcoming DESI (Schlegel et al. 2011; Levi et al. 2013; DESI Collaboration et al. 2016) survey, Large Synoptic Survey Telescope (LSST; LSST Science Collaboration et al. 2009) surveys, *Euclid* satellite mission surveys (Laureijs et al. 2011), and *WFIRST* (Green et al. 2012) surveys. Monaco (2016) estimates that for these future surveys, generating 1000 mock realizations with the fastest of the popular mock codes currently in use would take $\sim 1\,000\,000$ CPU hours since, unfortunately, the cheapest mock catalogues to produce, the lognormal mocks (Coles & Jones 1991; Beutler et al. 2011; Pearson, Samushia & Gagrani 2016), do not adequately reproduce the three-point statistics (see White, Tinker & McBride 2014, figure 9).

It would also be interesting to test if various methods of reducing the uncertainties in the covariance matrix such as shrinkage estimation (Pope & Szapudi 2008), estimation from fitting formula (Pearson & Samushia 2016), or calculating the expected covariance from theory (Xu et al. 2012), could work for the bispectrum. Slepian et al. (2017a,b) used a hybrid approach, fitting a theoretical model to the sample covariance from a 299 mocks, which may also be a useful approach for the bispectrum.

Lastly, although we find our theoretical template to be unbiased for this analysis, it would be interesting to test if using a more complex bispectrum model, such as the one used by Gil-Marín et al. (2015) and Gil-Marín et al. (2017), would affect the constraints. We leave these matters for exploration in future works.

ACKNOWLEDGEMENTS

LS is grateful for the support from Swiss National Science Foundation grant SCOPES IZ73Z0 152581, Georgia National Science Foundation grant FR/339/6-350/14, National Aeronautics and Space Administration (NASA) grant 12-EUCLID11-0004, and the Department of Energy grant DE-SC0011840. NASA’s Astrophysics Data System Bibliographic Service and the arXiv e-print service were used for this work. Additionally, we wish to acknowledge GNUPLLOT, a free open-source plotting utility that was used to create all of our figures.

REFERENCES

- Alam S. et al., 2017, *MNRAS*, 470, 2617
- Alcock C., Paczynski B., 1979, *Nature*, 281, 358
- Anderson L. et al., 2012, *MNRAS*, 427, 3435
- Anderson L. et al., 2014, *MNRAS*, 441, 24
- Baldauf T., Mercolli L., Mirbabayi M., Pajer E., 2015, *J. Cosmol. Astropart. Phys.*, 5, 007
- Ballinger W. E., Peacock J. A., Heavens A. F., 1996, *MNRAS*, 282, 877
- Beutler F. et al., 2011, *MNRAS*, 416, 3017
- Beutler F. et al., 2017, *MNRAS*, 464, 3409
- Coles P., Jones B., 1991, *MNRAS*, 248, 1
- Cuesta A. J. et al., 2016, *MNRAS*, 457, 1770
- Dawson K. S. et al., 2013, *AJ*, 145, 10
- DESI Collaboration et al., 2016, preprint (arXiv:1611.00036)
- Eisenstein D. J., Hu W., 1998, *ApJ*, 496, 605
- Eisenstein D. J., Seo H.-J., Sirko E., Spergel D. N., 2007, *ApJ*, 664, 675
- Feldman H. A., Kaiser N., Peacock J. A., 1994, *ApJ*, 426, 23
- Gagrani P., Samushia L., 2017, *MNRAS*, 467, 928
- Gil-Marín H., Noreña J., Verde L., Percival W. J., Wagner C., Manera M., Schneider D. P., 2015, *MNRAS*, 451, 539
- Gil-Marín H. et al., 2016, *MNRAS*, 460, 4210
- Gil-Marín H., Percival W. J., Verde L., Brownstein J. R., Chuang C.-H., Kitaura F.-S., Rodríguez-Torres S. A., Olmstead M. D., 2017, *MNRAS*, 465, 1757

- Green J. et al., 2012, preprint ([arXiv:1208.4012](https://arxiv.org/abs/1208.4012))
- Hartlap J., Simon P., Schneider P., 2007, *A&A*, 464, 399
- Hastings W. K., 1970, *Biometrika*, 57, 97
- Hu W., Sugiyama N., 1996, *ApJ*, 471, 542
- Jeong D., 2010, PhD thesis, University of Texas at Austin
- Kaiser N., 1987, *MNRAS*, 227, 1
- Kitaura F.-S. et al., 2016, *MNRAS*, 456, 4156
- Laureijs R. et al., 2011, preprint ([arXiv:1110.3193](https://arxiv.org/abs/1110.3193))
- Levi M. et al., 2013, preprint ([arXiv:1308.0847](https://arxiv.org/abs/1308.0847))
- Lewis A., Challinor A., Lasenby A., 2000, *ApJ*, 538, 473
- LSST Science Collaboration et al., 2009, preprint ([arXiv:0912.0201](https://arxiv.org/abs/0912.0201))
- Martínez V. J., Saar E., 2002, *Statistics of the Galaxy Distribution*, Chapman & Hall/CRC
- Monaco P., 2016, *Galaxies*, 4, 53
- Pearson D. W., Samushia L., 2016, *MNRAS*, 457, 993
- Pearson D. W., Samushia L., Gagrani P., 2016, *MNRAS*, 463, 2708
- Peebles P. J. E., 1980, *The large-scale structure of the universe*, Princeton University Press
- Percival W. J. et al., 2014, *MNRAS*, 439, 2531
- Planck Collaboration et al., 2016, *A&A*, 594, A13
- Pope A. C., Szapudi I., 2008, *MNRAS*, 389, 766
- Rodríguez-Torres S. A. et al., 2016, *MNRAS*, 460, 1173
- Ross A. J. et al., 2017, *MNRAS*, 464, 1168
- Saito S., Baldauf T., Vlah Z., Seljak U., Okumura T., McDonald P., 2014, *Phys. Rev. D*, 90, 123522
- Sakharnykh N., 2015, GPU Pro Tip: Fast Histograms Using Shared Atomics on Maxwell, <https://devblogs.nvidia.com/parallelforall/gpu-pro-tip-fast-histograms-using-shared-memory-atomics-maxwell/> (Accessed on June 14, 2018)
- Samushia L. et al., 2011, *MNRAS*, 410, 1993
- Samushia L., Branchini E., Percival W. J., 2015, *MNRAS*, 452, 3704
- Schlegel D. et al., 2011, preprint ([arXiv:1106.1706](https://arxiv.org/abs/1106.1706))
- Scoccimarro R., 2000, *ApJ*, 544, 597
- Scoccimarro R., 2015, *Phys. Rev. D*, 92, 083532
- Scoccimarro R., Feldman H. A., Fry J. N., Frieman J. A., 2001, *ApJ*, 546, 652
- Sefusatti E., Scoccimarro R., 2005, *Phys. Rev. D*, 71, 063001
- Simpson F., Peacock J. A., 2010, *Phys. Rev. D*, 81, 043512
- Slepian Z., Eisenstein D. J., 2015, *MNRAS*, 454, 4142
- Slepian Z. et al., 2017a, *MNRAS*, 468, 1070
- Slepian Z. et al., 2017b, *MNRAS*, 469, 1738
- Spergel D. et al., 2015, preprint ([arXiv:1503.03757](https://arxiv.org/abs/1503.03757))
- Taylor A., Joachimi B., Kitching T., 2013, *MNRAS*, 432, 1928
- Thuerck D., Widmer S., Kuijper A., Goesele M., 2014, in Wyrzykowski R., Dongarra J., Karczewski K., Wańniewski J., eds., *Parallel Processing and Applied Mathematics*, Springer, Berlin, Heidelberg, p. 652
- White M., Tinker J. L., McBride C. K., 2014, *MNRAS*, 437, 2594
- Xu X., Padmanabhan N., Eisenstein D. J., Mehta K. T., Cuesta A. J., 2012, *MNRAS*, 427, 2146

APPENDIX A: THE TRIANGLE INDEX EXPLAINED

The spherically averaged bispectrum is a function of three⁶ independent variables. This makes plotting the bispectrum difficult. One possibility is to compress the information in some way by integrating over one of the k 's in the triplet as we did in Fig. 3. However, in doing so we are, by definition, getting rid of some information.

The other option is to assign each k -triplet bin a unique identifier, *the triangle index*. There really is no physical meaning to this number, it is simply the element number of the particular bispectrum estimate in our data vector. The triangle index 1 starts with the lowest value of k considered for each of the three values or, 0.044, 0.044, and 0.044. The triangle index 2 simply adds the bin width to the value of k_3 yielding 0.044, 0.044, and 0.052. This repeats until the triangle condition is violated, at which point k_2 is incremented and k_3 rolls back to be equal to k_2 , and we again begin incrementing k_3 until the triangle condition is violated, or we reach the maximum k being considered. Once k_2 hits the maximum, we increment k_1 and repeat the whole process again.

This paper has been typeset from a \LaTeX file prepared by the author.

⁶Really only two as the third is fixed by the triangle condition. However, binning the centers of the bins may not exactly solve the triangle condition which is why we explain it as being three independent variables